

Quellenangaben und Links zum Artikel

„Systematische Zusammenhänge und typische Umsetzungsfehler: Leistungsbeurteilungen gestalten“

von Prof. Dr. Boris Kaehler

(Arbeit und Arbeitsrecht Heft 11/2014)

Die Ausführungen im Beitrag basieren auf dem Buch „Komplementäre Führung – Ein praxiserprobtes Modell der organisationalen Führung“.

Kritik und Anregungen richten Sie gern an:
boris.kaehler@goodhr.de; www.goodhr.de

Exkurs A: 360°-Beurteilungen

Die sog. „360°-Beurteilung“ (auch: „360°-Feedback“) beinhaltet die Beurteilung von mehreren Seiten gleichzeitig, wobei neben der Führungskraft, den Kollegen und ggf. den eigenen Mitarbeitern teilweise auch Kunden, Lieferanten, Top-Management sowie Selbstbeurteilungen einbezogen werden. Eine ebenso häufige wie fatale Fehlkonstruktion besteht darin, diese Einschätzungen nicht – wie im Artikel empfohlen – als Ausgangsdaten zu verstehen, sondern als vollwertige Beurteilung gelten zu lassen. Die jeweiligen Personengruppen können ja jeweils nur einen sehr begrenzten Ausschnitt der Gesamtleistung einschätzen und tun dies höchst subjektiv. Das kann zu völlig verzerrten Beurteilungen führen und fatale Folgen für den Leistungsprozess und die Motivation des Beurteilten haben. Zudem ergeben sich Probleme daraus, dass sich Stelleninhaber gegenseitig beurteilen und somit Gefälligkeits- und Rachebeurteilungen vorkommen können.

Auf Grund der mangelnden Objektivität und vielfältigen methodischen Probleme der Teilerückmeldungen nutzen viele Organisationen 360°-Verfahren bewusst nicht als Beurteilungsinstrumente, sondern lediglich als Grundlage informeller Feedbacks; die Datenaufbereitung und Ergebniserückmeldung wird dabei i.d.R. zwecks Vertraulichkeit an externe Dienstleister vergeben. Dieses Vorgehen ist insofern positiv zu sehen, als es immerhin ein umfassendes und differenziertes Feedback gewährleistet und für den Beurteilten wichtige Motivations- und Handlungsimpulse beinhaltet. Allerdings ist dies durch eine Entwertung der eigentlichen Beurteilungen erkauft. Fortan verfügt nur noch der Mitarbeiter selbst über alle wesentlichen Informationen, während die übrigen Führungsakteure (Führungskraft, obere Führungskraft, Personalbetreuer) nur noch auf einen kleinen Teil der leistungsrelevanten Informationen zugreifen können. Dadurch können dysfunktionale Parallelsysteme entstehen: Je stärker der Impuls aus dem Feedback, desto eher verdrängt er Impulse aus der Leistungsbeurteilung.

Beispiel:

Produktentwickler Huber erhält eine positive Leistungsbeurteilung seiner Führungskraft und, aus dem vertraulichen 360°-Feedback, sehr negative Rückmeldungen aus der Marketingabteilung (die teilweise andere Ziele verfolgt). Als harmonieorientierter Mensch wird er sich fortan bemühen, die Zufriedenheit der Kollegen zu steigern und dafür ggf. Kompromisse zu Lasten der eigenen Arbeit in Kauf zu nehmen.

Exkurs B: Komplementäres Zusammenwirken der Führungsakteure

Leistungsbeurteilung und Leistungsfeedback sind Führungsaufgaben. Folgt man den Annahmen des Komplementären Führungsmodells (AuA 8/2014), so obliegen Führungsaufgaben jedoch nicht in erster Linie der Führungskraft, sondern dem Mitarbeiter selbst, der dabei durch andere Führungsakteure unterstützt wird. So auch hier: Der Mitarbeiter sollte seine Leistung selbst beurteilen, also sich die leistungsrelevanten Informationen selbst verschaffen und diese selbst beurteilen. Nur falls dies nicht geschieht, greifen die Führungskraft, die Kollegen, die obere Führungskraft und/oder die Personalfunktion kompensatorisch ein. Formalisierte Beurteilungsroutinen und -instrumente können diesen Prozess lenken und unterstützen.

Beispiel:

Verkäufer Müller hat laufend Zugang zu seinen wesentlichen Leistungsdaten (z.B. Umsatzkennziffern und Kundenbewertungen) und holt sich bei Bedarf ergänzende Informationen ein (z.B. die Kundenhistorie oder die Einschätzung von Kollegen). Im wöchentlichen Mitarbeitergespräch oder im Jahresbeurteilungsgespräch gibt er – in einer Art eines Fachgesprächs – zunächst seine Eigenbeurteilung ab, die von der Führungskraft ggf. ergänzt oder korrigiert wird.

Was die Führungsaufgabe „Leistung beurteilen und Feedback geben“ auf den ersten Blick verwirrend macht, ist die Tatsache, dass nicht nur die abschließende Beurteilung von unterschiedlichen Führungsakteuren übernommen werden kann (Schritt 2), sondern dass die Leistungseinschätzung aller Akteure auch schon in der Phase der Datenerhebung (Schritt 1) auch als Ausgangs- bzw. Rohdatum einfließen kann. Auch dort wirken zwar unterschiedliche Führungsakteure zusammen, aber nicht mit dem Anspruch einer ganzheitlichen Würdigung der leistungsbezogenen Datenlage.

Exkurs C: Methodische Probleme und entsprechende Lösungsansätze

Im Artikel wurde bereits deutlich, dass das Einbeziehen vielfältiger Leistungsdaten dem Bemühen um Validität (i.S.d. Abbildung der tatsächlichen Arbeitsleistung) geschuldet ist. Gleiches gilt für die Art der Bewertung: Die Komplexität der Interpretation anhand flexibler Bewertungsmaßstäbe muss in Kauf genommen werden, um reale Arbeitsleistung und nicht etwa nur ein Teil- oder Zerrbild zu erfassen. Messmethodisch allerdings ist dies alles andere als optimal. Es ergeben sich vier potenzielle methodische Probleme.

Bewertungsgenauigkeit: Je mehr Daten einbezogen werden und je komplexer der Bewertungsprozess, desto größer die Fehleranfälligkeit. Dabei ist es hilfreich, wenn jedes Beurteilungskriterium durch mehrere Ausgangsdatenarten erfasst wird. Auch der Umstand, dass die Beurteilung nicht nur jährlich durchgeführt wird, sondern im Rahmen regelmäßiger Mitarbeitergespräche wiederholt in kürzeren Abständen, kommt der Genauigkeit zu Gute. Ferner müssen klassische Beurteilungsfehler thematisiert werden; diese sind freilich überwiegend gar nicht spezifisch für die Leistungsbeurteilung, sondern stellen eher allgemeine Denk- und Entscheidungsfehler dar, deren Vermeidung schon im Interesse einer guten Sacharbeit trainiert werden sollte.

Objektivität: Um trotz der freien Interpretation und der flexiblen Beurteilungsmaßstäbe/Soll-Werte ein Mindestmaß an Objektivität (im Sinne von Beurteilerunabhängigkeit) zu erreichen, sind klare Entscheidungsvorgaben, flächendeckende Beurteilungstrainings und ein diskursiver Abgleich der Beurteilungsergebnisse erforderlich. Zudem muss im Bewertungsprozess die subjektive Natur der Ausgangsdaten berücksichtigt werden, was Informationen über ihr Zustandekommen voraussetzt.

Transparenz: Auch die Transparenz für die Beurteilten lässt sich weitgehend retten. Zwar kann nur derjenige, der weiß, woran er gemessen wird, sein eigenes Verhalten entsprechend steuern und flexible Bewertungsmaßstäbe können dies unterminieren. Dem lässt sich jedoch vorbeugen, indem die mit der erwarteten Leistung intendierten Zielvorstellungen sowie mögliche Interpretationsspielräume eingehend besprochen werden und die spätere Bewertung dem Geist der ursprünglichen Leistungsidee treu bleibt. Ferner basiert das hier empfohlene Modell auf einer Kombination von Jahres- und Wochenbeurteilungen. Der Mitarbeiter erfährt also in kurzen Abständen von evtl. veränderten Soll-Vorstellungen und kann seine Leistung entsprechend anpassen.

Missbrauch: Schließlich muss auch der Missbrauch der Beurteilungsfreiheiten zu mikropolitischen Zwecken nicht hingenommen werden. Natürlich sind Beurteilende nie frei von Eigeninteressen und verfolgen mit ihrer Bewertung u.U. mikropolitische Strategien. Dies ist auf ein erträgliches Maß zu reduzieren, wenn mehrere Akteure am Beurteilungsprozess mitwirken und sich gegenseitig kontrollieren (die berühmten „Checks and Balances“). Das vom Autor in AuA 8/2014 vorgestellte Komplementäre Führungsmodell bietet einen Rahmen dafür.

Exkurs D: Ergebnisverteilung und Leistungsdifferenzierung

Eine breite Streuung der Beurteilungsergebnisse innerhalb der beurteilten Mitarbeiterschaft wird vor allem deshalb als wünschenswert betrachtet, weil diese die Ableitung von Personalentscheidungen erleichtert. Wissenschaftlicherseits wird zudem vertreten, dass differenzierte Ergebnisse mit höherer Leistungsbereitschaft und Produktivität einhergehen. Um eine breitere Streuung zu erreichen, sind nun grundsätzlich zwei Wege gangbar.

Wo Leistung anhand absoluter oder individueller Maßstäbe bewertet wird, lassen sich schlicht die Skalen an das erfahrungsgemäß mögliche Leistungsniveau anpassen. Absolute Zielwerte legen ja feste quantitative oder qualitative Werte fest (z.B. Fehlerfreiheit; Umsatz >X), individuelle Zielwerte solche in Bezug auf individuelle Maßstäbe (z.B. weniger Fehler und mehr Umsatz als eigene Vorjahreswerte). In beiden Fällen aber fließen in die Festlegung zwangsläufig auch Überlegungen in Bezug auf das von anderen Erreichte und damit prinzipiell Erreichbare ein. Eine Anpassung der Skala bzw. der Zielwerte an diese Erfahrungswerte führt – Beurteilungsfehler und wünschenswerte Motivationseffekte seien hier einmal vernachlässigt – zu einer ausgeprägteren Differenzierung.

Beispiel:

Alle Teammitglieder übertreffen oder verfehlen ihre Umsatzziele bei weitem. Die Ziele werden nun genau so viel höher gesetzt, dass im nächsten Durchgang nur noch die besten sehr gut beurteilt werden.

Der zweite Weg ist der einfachere und besteht darin, solche relativen Bewertungsmaßstäbe anzuwenden, die mittels Ranglistenbildung („Ranking“) einen vergleichenden Bezug innerhalb der Gruppe herstellen. Insbesondere die Rangbildung unter Vorgabe fester Ergebnisquoten („Forced Distribution“) sorgt in in Fachkreisen seit Jahren für Diskussionen.

Beispiel:

Der Konzern Unilever lässt seine Belegschaft in vier Leistungs- und Potenzialkategorien einteilen, denen unterschiedliche Farben zugeordnet sind (z.B. rot für Minderleistung). Bei außertariflichen Mitarbeitern ist dabei eine Normalverteilung vorgeschrieben, sodass automatisch ein fester Anteil der Beurteilten die Kategorie rot erhält. Bei General Electrics wurde langjährig das 20/70/10-Modell praktiziert, wonach die besten 20% mit Belohnungen überschüttet, die mittleren 70% gefördert und die schlechtesten 10% gekündigt werden.

In Literatur und Praxis werden u.a. die folgenden Nachteile solcher Verfahren herausgestellt:

- Die erzwungene Gesamtverteilung entspricht nie der realen Verteilung.
- Der vergleichende Maßstab steht in logischem Widerspruch zu festen Zielvereinbarungen.
- Mikropolitische Überlegungen treten in den Vordergrund („wer braucht welche Bewertung am Dringenden?“).
- Der interne Wettbewerb wird gefördert; dies untergräbt den Teamgedanken und kann Mitarbeiter zur gegenseitigen Sabotage verleiten.
- Das Verfahren wird von den Mitarbeitern häufig als unfair empfunden und ist unbeliebt.

Die Verwendung fester Verteilungsquoten ist daher abzulehnen. Je nach statistischer Grundgesamtheit kann es passieren, dass in einem Beurteilungsdurchgang auch viele Schlechtleister im oberen Segment landen, während im nächsten Beurteilungsdurchgang auch viele Hochleister schlecht eingestuft werden. Dies zerstört den Bezug der Beurteilung zur eigentlichen Arbeitsleistung und nimmt den Betroffenen jedes Kontrollbewusstsein – ob andere besser sind und sie aus einer Leistungskategorie verdrängen, können sie nicht beeinflussen.

In der Regel sind solche Quoten ein Indiz für Personalarbeit, die mehr durch administrative Vorgaben als durch echte Leistungsorientierung geprägt ist, denn sie treten auf, wo die Personalabteilung feste Maßnahmenkontingente zu vergeben hat. Dass aber nun genau X % der Mitarbeiter sich weiterbilden dürfen, einen Bonus erhalten oder gar – wie bei General Electrics – gehen müssen, kann schwerlich bedarfsgerecht sein. Wohlbemerkt: Das Prinzip der Leistungsdifferenzierung, wonach Minder- und Topleistung klar herausgearbeitet und mit entsprechenden Personalmaßnahmen verknüpft werden, ist uneingeschränkt zu begrüßen – nur eben nicht nach starren Quoten.

Dagegen kann Rangfolgenbildung ohne Quoten durchaus sinnvoll sein, allerdings nur in der Phase der Ausgangsdatenerhebung (Schritt 1). Steht ein Mitarbeiter am unteren Ende der Umsatzliste und wird von Kunden/Kollegen deutlich hilfreicher eingeschätzt als andere, so sind dies wichtige Leistungsaspekte, gegen deren Einbezug kaum etwas einzuwenden ist. Auf Ebene der Bewertung (Schritt 2) sieht dies anders aus. Hier darf es nicht bei einer vergleichenden Einordnung bleiben, sondern es sind ihre Implikationen zu bewerten: Bedeutet der schlechte Rang – was natürlich durchaus der Fall sein kann – auch eine schlechte Leistung?

Quellenangaben

Abschnitt 1: Beurteilungen als Führungsaufgabe und Personalinstrument

- **Malik-Zitat:** Malik, Fredmund (2006): „Führen, Leisten, Leben“, Neuauflage Campus 2006 (Erstauflage 2000); S. 348ff.
- **Fundamentale HR-Kritik** ohne Schwerpunkt auf Beurteilungen, sie aber verschiedentlich aufgreifend: Spilker, Martin/Roehl, Heiko/Hollmann, Detlef (2013): „Die Akte Personal – Warum sich die Personalwirtschaft jetzt neu erfinden sollte“; Verlag Bertelsmann Stiftung 2013.
- **Überblick über die wesentlichen Aspekte, Verfahren und Methoden der Personalbeurteilung:** z.B. bei Stock-Homburg, Ruth (2013): „Personalmanagement – Theorien, Konzepte, Instrumente“; 3. Auflage Springer-Gabler 2013; S. 359-398; Berthel, Jürgen/Becker, Fred G. (2013): „Personalmanagement – Grundzüge für Konzeptionen betrieblicher Personalarbeit“; 10. Auflage Schäffer Poeschel 2013; S. 269-298; Blickle, Gerhard (2014): „Leistungsbeurteilung“; in Nerdinger, Friedemann W./Blickle, Gerhard/Schaper, Niclas: „Arbeits- und Organisationspsychologie“; 3. Auflage Springer 2014; S. 271-289; Lohaus, Daniela/Schuler, Heinz (2014): „Leistungsbeurteilung“; in Schuler, Heinz/Kanning, Uwe Peter: „Lehrbuch der Personalpsychologie“; 3. Auflage Hogrefe 2014; S. 357-411.

Abschnitt 2: Warum beurteilen? Zwei Funktionen

- Die **Unterscheidung von Feedback- und Administrationsfunktion** ist durchaus üblich: vgl. z.B. Blickle, Gerhard (2014): „Leistungsbeurteilung“; in Nerdinger, Friedemann W./Blickle, Gerhard/Schaper, Niclas: „Arbeits- und Organisationspsychologie“; 3. Auflage Springer 2014; S. 272f., der noch eine Forschungsfunktion ergänzt.
- Auf die **Wichtigkeit des Leistungsfeedbacks für die Selbststeuerung** und damit für die Steuerung der Arbeitsleistungen insgesamt hat vor allem Peter Drucker immer wieder hingewiesen und diesen Aspekt in den Mittelpunkt seines Konzepts „Führen durch Ziele und Selbstkontrolle“ gestellt: Drucker, Peter F. „Was ist Management - Das Beste aus 50 Jahren“, Ullstein Buchverlage, 5. Auflage 2007, S. 150f.
- Mehr zum **Thema Motivation:** „Arbeitsmotivation – Ein integratives Prozessmodell“ (Boris Kaehler); Arbeit und Arbeitsrecht 5/2014; S. 288-291.

Abschnitt 3: Alle Beurteilungsanlässe und -zeiträume einbeziehen

- **Regelbeurteilungen und informelle Leistungsrückmeldungen als Teile eines Systems zu behandeln**, liegt auf der Hand: vgl. z.B. Blickle, Gerhard (2014): „Leistungsbeurteilung“; in Nerdinger, Friedemann W./Blickle, Gerhard/Schaper, Niclas: „Arbeits- und Organisationspsychologie“; 3. Auflage Springer 2014; S. 271-289; Dulisch, Frank (2014): „Grundlagen der Personalbeurteilung“; in Sauer, Joachim/Cisik, Alexander: „In Deutschland führen die Falschen – Wie sich Unternehmen ändern müssen“; Bundesverband der Personalmanager e.V. 2014; S. 170-191.

Abschnitt 4: Beurteilungsgegenstand und Beurteilungskriterien – Was bedeutet „Leistung“?

- Kritisch zur herrschenden Philosophie rein **kennzifferngesteuerter Zielsysteme**: Claßen, Martin/Sattelberger, Thomas (2011): „Vor dem Platzen der »Leadership Bubble«“; Organisationsentwicklung 2/2011; S. 61.
- Die angeführte Entscheidung des **Bundesarbeitsgerichts**: BAG 11.12.2003; 2 AZR 667/02.

Abschnitt 5: Generieren der Ausgangsdaten – Zwei Erhebungsvorgänge

- **Überblick über die wesentlichen Aspekte und methodischen Probleme unterschiedlicher Datenerhebungs- bzw. Datenerfassungsverfahren**: z.B. bei Stock-Homburg, Ruth (2013): „Personalmanagement – Theorien, Konzepte, Instrumente“; 3. Auflage Springer-Gabler 2013; S. 359-398; Berthel, Jürgen/Becker, Fred G. (2013): „Personalmanagement – Grundzüge für Konzeptionen betrieblicher Personalarbeit“; 10. Auflage Schäffer Poeschel 2013; S. 269-298; Dulisch, Frank (2014): „Grundlagen der Personalbeurteilung“; in Sauer, Joachim/Cisik, Alexander: „In Deutschland führen die Falschen – Wie sich Unternehmen ändern müssen“; Bundesverband der Personalmanager e.V. 2014; S. 170-191; Blickle, Gerhard (2014): „Leistungsbeurteilung“; in Nerdinger, Friedemann W./Blickle, Gerhard/Schaper, Niclas: „Arbeits- und Organisationspsychologie“; 3. Auflage Springer 2014; S. 271-289.

Abschnitt 6: Abwägen statt Abzählen: Die eigentliche Beurteilung

- **Drucker-Zitat**: Drucker, Peter F. „Was ist Management - Das Beste aus 50 Jahren“, Ullstein Buchverlage, 5. Auflage 2007, S. 152.
- Zum Thema **Subjektivität**: Kampkötter, Patrick/Sliwka, Dirk (2013): „Trotz subjektiver Leistungsbeurteilung zu aussagekräftigen Ergebnissen kommen“; Personal Quarterly 2/2013; S. 46-49.
- Zum Missbrauch der Beurteilung zu **mikropolitischen Zwecken**: vgl. z.B. Dulisch, Frank (2014): „Grundlagen der Personalbeurteilung“; in Sauer, Joachim/Cisik, Alexander: „In Deutschland führen die Falschen – Wie sich Unternehmen ändern müssen“; Bundesverband der Personalmanager e.V. 2014; S. 181f.
- Zum **Komplementären Führungsmodell**: Kaehler, Boris: „Komplementäre Führung - Mitarbeiterführung wirksam gestalten“; Arbeit und Arbeitsrecht 8/2014; S. 459-461.

Abschnitt 9: Rückmeldung und Nutzung der Ergebnisse

- Zum **Selbstwertbezug aller Kommunikation**: vgl. z.B. Birkenbiel, Vera F.: „Kommunikationstraining“; 27. Auflage 2006; S. 21ff.

Exkurs A: 360°-Beurteilungen

- Zu **360°-Beurteilungen** und ihren Problemen: Heiden, Thomas K. (2011): „360°-Feedback“; Arbeit und Arbeitsrecht 8/2011; S. 475-477; Groth, Julia (2012): „Rundum beäugt“; Human Resources Manager Juni/Juli 2012; S. 55-57; Dembkowski, Sabine/Althaus, Ulrich (2013): „Fingerspitzengefühl gefragt“; Personalwirtschaft 7/2013; S. 42-44; Scherm, Martin (2013): „360-Grad-Beurteilungen“ in Sarges, Werner „Management-Diagnostik“; 4. Auflage Hogrefe 2013; S. 864-872.

Exkurs B: Komplementäres Zusammenwirken der Führungsakteure

- Zum **Komplementären Führungsmodell**: Kaehler, Boris: „Komplementäre Führung - Mitarbeiterführung wirksam gestalten“; Arbeit und Arbeitsrecht 8/2014; S. 459-461; Kaehler, Boris (2014): „Komplementäre Führung – Ein praxiserprobtes Modell der organisationalen Führung“; Springer Gabler 2014.

Exkurs C: Methodische Probleme und entsprechende Lösungsansätze

- Zu **methodischen Problemen von Beurteilungen Datenerhebungs- bzw. Datenerfassungsverfahren**: vgl. z.B. Blickle, Gerhard (2014): „Leistungsbeurteilung“; in Nerdinger, Friedemann W./Blickle, Gerhard/Schaper, Niclas: „Arbeits- und Organisationspsychologie“; 3. Auflage Springer 2014; S. 271-289; Lohaus, Daniela/Schuler, Heinz (2014): „Leistungsbeurteilung“; in Schuler, Heinz/Kanning, Uwe Peter: „Lehrbuch der Personalpsychologie“; 3. Auflage Hogrefe 2014; S. 357-411.

Exkurs D: Ergebnisverteilung und Leistungsdifferenzierung

- Zum **Einfluss differenzierter Ergebnisse** auf Leistungsbereitschaft und Produktivität: vgl. z.B. Berger, Johannes/Harbring, Christine/Sliwka, Dirk (2013): „Performance Appraisals and the Impact of Forced Distribution: An Experimental Investigation“; Management Science January 2013; S. 54-68.
- Zum Prinzip der **erzwungener Verteilung („Forced Distribution“)**: vgl. z.B. Weilbacher, Jan (2013): „Zum Glück gezwungen“; Human Resources Manager Juni/Juli 2013; S. 64-65; Dulisch, Frank (2014): „Grundlagen der Personalbeurteilung“; in Sauer, Joachim/Cisik, Alexander: „In Deutschland führen die Falschen – Wie sich Unternehmen ändern müssen“; Bundesverband der Personalmanager e.V. 2014; S. 188f.; Budras, Corinna (2014): „Wie die Firmen ihre guten Leute demotivieren“; Frankfurter Allgemeine Sonntagszeitung 1.6.2014; S. 25.
- Zum **Beurteilungssystem bei Unilever**: Weilbacher, Jan (2013): „Wir wollen eine Leistungskultur“ (Interview mit Ulf Werkmeister); Human Resources Manager Februar/März 2013; S. 53-55
- Zum **20/70/10-Modell bei General Electrics**: Welch, Jack/Welch, Suzy (2005): „Winning – Das ist Management“; Campus 2005; S. 52ff.